



## Background & Motivations

Police departments are increasingly relying on **algorithms to predict where future crime will occur**.<sup>1</sup> These algorithms are flawed in that they:

- Are used by departments to predict crimes they aren't designed to predict.<sup>2</sup>
- Treat all crimes as equal in dispersion patterns.
- **Lack sensitivity to racial/socioeconomic bias**.<sup>3</sup>

These deficits come both from **bad training data** and the inherent functioning of these algorithms.

Our team aims to **create a new general-purpose predictive algorithm** which not only better predicts crime, but **accounts for historical inequities** in crime enforcement.

## Research Question

Is it possible to make an algorithm which is as or **more accurate** than current prediction software, but **less biased**?

## Data Collection

Crime data in **Montgomery County, MD** (300k crimes)<sup>4</sup> and **Chicago, IL** (800k crimes)<sup>5</sup>

FOIA requests: **Charles County Police Department, MD; DC Metro Police; and Prince George's County Police Department, MD.**

## Further Goals & Research

While our method is more modern than previous attempts at predictive policing, the algorithm needs to be designed with unfairness in mind. To that end, we are working to implement demographic data that will allow the algorithm to be **penalized for unfairness**.<sup>6</sup>

We aim to create a website that allows departments to visually analyze crime data on a map. They will also be able to input their own crime data, contributing to the overall project of training the algorithm.

## References

Special thanks to our mentor Mohammad Hajiaghayi and our librarian Celina McDonald.



## How it Works

### Historical Bias

Algorithms rely on data that reflect **historical prejudice**.

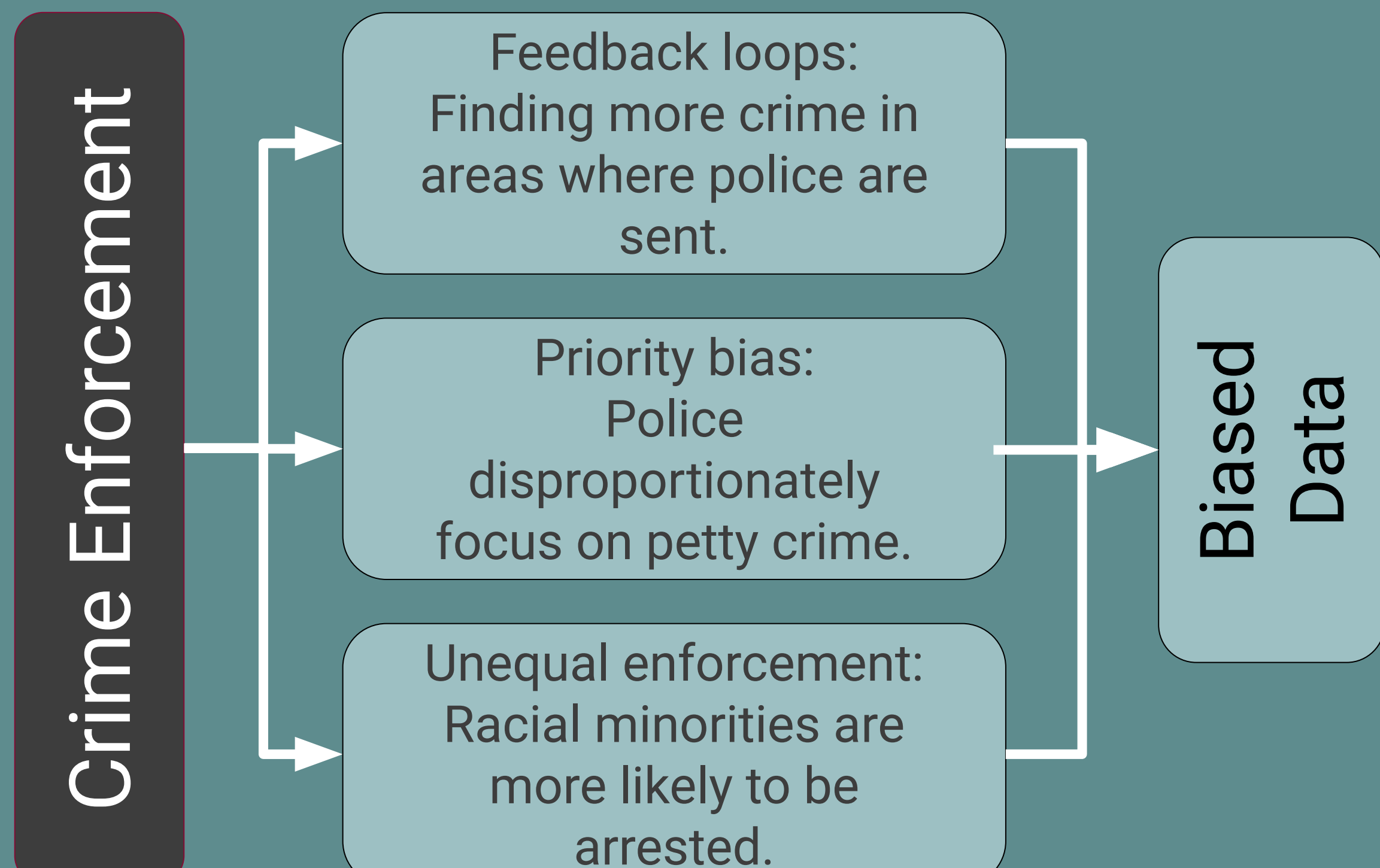


Figure 1: How Crime Data Becomes Biased

Divide into grid & label according to crime type

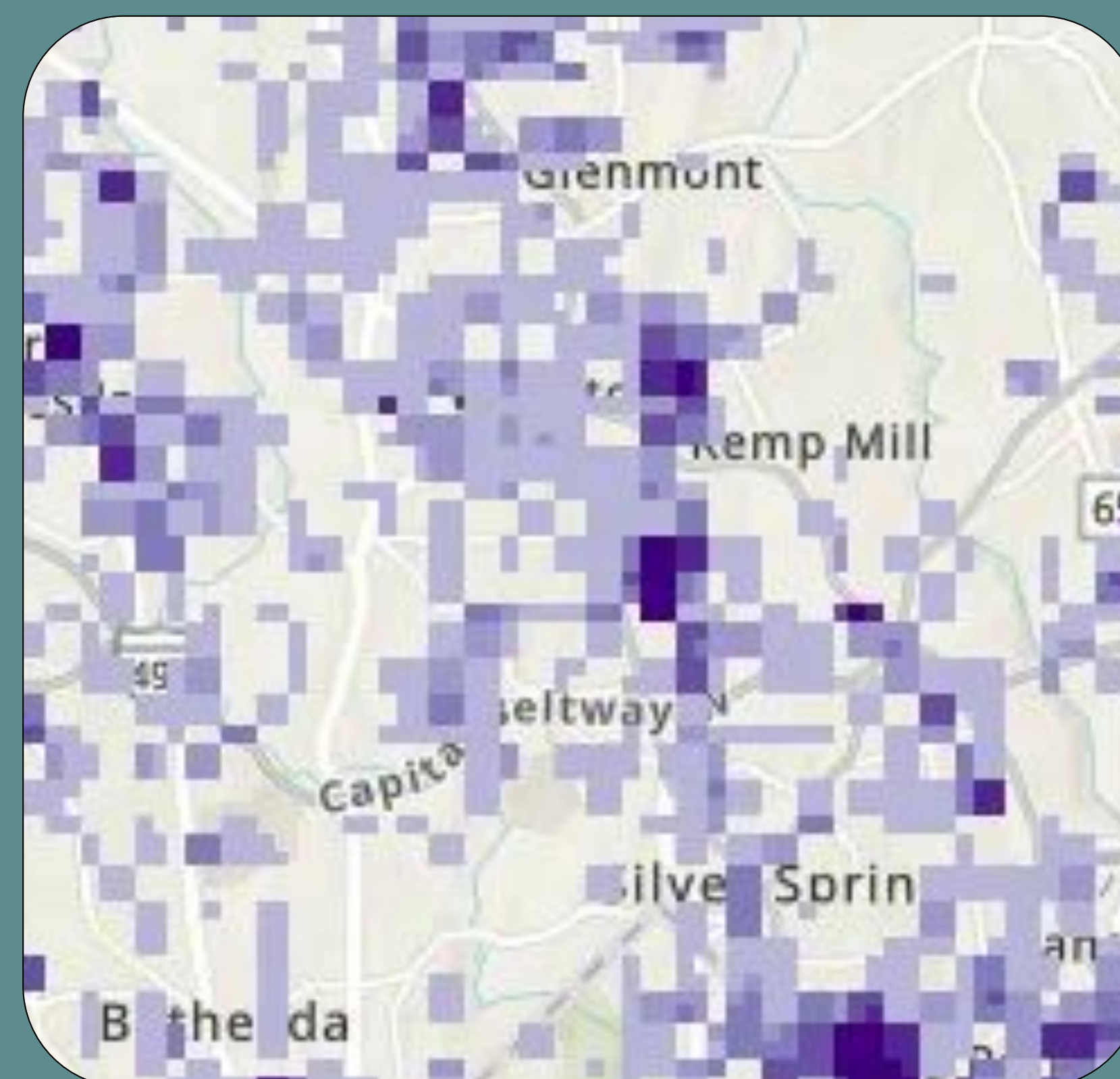


Figure 2: Montgomery County Crime Heat Map

Use AutoML to run and rank multiple models

Model Name	Score
Temporal Fusion Transformer	0.975
Weighted Ensemble	0.977
ARIMA	2.619
Seasonal Naive	11.36

Incorporate a fairness function to penalize unequal policing of minorities, then retrain<sup>6</sup>

Run highest ranked model to create hot spots